

PERCEPCION DE LA ASPEREZA DE SEÑALES ACUSTICAS URBANAS

PACS: 43.66.Lj

A. Moreno, F. Simón, C. de la Colina
Instituto de Acústica del CSIC
Serrano 144
28006 Madrid. España
E-Mail: iacma15@pinar2.csic.es

ABSTRACT

A thorough study of main processes involved in roughness perception according to the model of Terhardt-Aures-Daniel (TAD) is presented, with emphasis on aspect regarding sound quality of environmental urban noise.

Generalised modulation factor is considered the capital factor involved in roughness and a new method for extraction of this factor is proposed and compared to that of Terhardt. It is shown that excitation signals regarded at the critical bands level are true amplitude modulated signals then able to support computings of the modulation factor.

Mainly based on the loudness detection threshold, a rejection operator has been constructed to discard contributions to total roughness of critical bands whose specific excitation levels are below the threshold.

As conclusion divergences among direct subjective evaluations and objective computed roughness of 32 environmental noises are commented and the possibility of a new concept of roughness is suggested.

RESUMEN

En este trabajo se profundiza en el conocimiento de los procesos fundamentales implicados en la percepción subjetiva de la aspereza con el fin de aportar conocimiento sobre su percepción en señales más complejas, en particular en los ruidos del entorno urbano, y cuantificar su contribución a la calidad acústica ambiental.

Se resalta aún más, si cabe, la importancia crucial del factor generalizado de modulación en el modelo TAD de valoración objetiva de la aspereza subjetiva, proponiendo un nuevo método que se

compara con el de Terhardt-Aures. Se fundamenta la detección subjetiva de este factor de modulación, clave en la percepción de este atributo, tanto para señales moduladas en amplitud como en señales moduladas en frecuencia. Para las señales moduladas en frecuencia se muestra que la validez del mismo se debe a que las señales de excitación específicas son verdaderas señales moduladas en amplitud por efecto del filtrado en cada una de las bandas subjetivas de frecuencia (bandas críticas).

Basándose en los conceptos de umbral de detección de sonoridad y en el umbral diferencial de sonoridad se justifica más sólidamente el rechazo de la contribución de las bandas críticas donde, a pesar de encontrar altos valores del factor de modulación generalizado, su contribución al efecto total es prácticamente nula.

Finalmente se plantean interrogantes sobre errores perceptivos de este atributo en sonidos complejos del ambiente acústico urbano, en donde hemos encontrado una gran divergencia entre los resultados de valoración subjetiva y los cálculos objetivos.

INTRODUCCION

En un estudio generalizado de la calidad perceptiva de señales acústicas del entorno urbano, emprendido conjuntamente entre un equipo de investigadores del Instituto de Acústica y otro del LMA (CNRS Marsella), encontramos discrepancias enormes entre los resultados de la aspereza percibida obtenida mediante un colectivo de sujetos [1] y la calculada mediante un programa desarrollado a partir del esquema de Daniel y Weber [2]. Un estudio amplio de estas discrepancias nos llevó a una profundización de los procesos implicados en la percepción de la aspereza, cuyas conclusiones más importantes constituyen el núcleo de esta contribución.

La aspereza, desde los tiempos de von Békésy, ha sido considerada como uno de los factores fundamentales negativos de la calidad perceptiva de señales acústicas. De hecho se incluye, junto con la sonoridad, la tonalidad y la agudeza, entre los cuatro factores significativos de la calidad de señales acústicas [3], [4]. En estos últimos años estamos asistiendo a un auge notable de los sistemas instrumentales de cálculo objetivo de la calidad percibida de señales, que en el caso particular del mundo del automóvil es verdaderamente espectacular [5]. En estos programas así como en otros usos es de importancia relevante disponer de algoritmos de cálculo que de manera objetiva obtengan resultados coincidentes con las valoraciones subjetivas, manteniendo la fiabilidad y reduciendo drásticamente el tiempo de obtención del resultado.

Hay dos percepciones de aspereza una en las señales de baja frecuencia y otra en altas frecuencias.

En un trabajo anterior [2] hemos señalado la escasa amplitud del contenido espectral en una banda crítica de frecuencia como un factor fundamental de detrimento del cálculo de la aspereza, particularmente en señales moduladas en amplitud. La similitud entre las señales de excitación en las diferentes bandas críticas de frecuencias (barks), debida al mecanismo subjetivo de migración espectral de las frecuencias contenidas en la señal a las bandas críticas, según las figuras patrón de migración (filtros alineales de Terhardt), lo que lleva a valores semejantes del factor generalizado de modulación, prácticamente independientes de la banda de frecuencia considerada. En ese trabajo proponíamos que las bandas críticas con contenidos energéticos 30 dB por debajo de las bandas más energéticas no contribuían a la aspereza percibida. Esta hipótesis que conducía a valores calculados en gran concordancia con resultados calculados publicados [6], tiene la debilidad de que el nivel de rechazo varía de unas señales a otras y se basa en el hecho conocido de que una señal con un nivel muy inferior a otra contribuye muy poco al nivel global resultante y supone que un paralelismo quizá excesivo entre la percepción de sonoridad y de aspereza. Esta necesidad de no contabilizar la aspereza parcial en ciertas bandas críticas falta en el esquema publicado del modelo

TAD, aunque se sospechaba el uso de algún mecanismo selectivo, que finalmente se ha descubierto semi-oculto en el texto.

MODELOS DE DETECCIÓN DEL GRADO DE MODULACION GENERALIZADO

La aspereza se considera un atributo perceptivo básico de la calidad de señales desde su mención por Helmholtz [7], que ha suscitado la atención de notables investigadores como el Nobel v. Békésy [8] o más recientes como Terhardt [9], Aures[10], Kemp [11], Vogel [12] y Plomp y Steeneken [13], sobre todo. También se ha resaltado su enorme influencia en la calidad de la conversación [14], de donde quizá tomó el nombre, que en nuestro idioma se asociaría tal vez mejor con ronquera, palabra ronca. Nosotros usaremos el término aspereza debido a que representando bien el concepto permite una derivación natural para su unidad de medida: el asperio. (Nótese que en inglés la unidad se denomina “asper”).

El concepto fundamental de la aspereza está ligado a los efectos de fluctuaciones de la señal, fluctuaciones de la amplitud en señales moduladas en amplitud. Pero el mismo atributo perceptivo se obtiene con señales moduladas en frecuencia. Las señales acústicas más típicas de la bibliografía con propiedades marcadas de aspereza, son tonos puros modulados en amplitud, cuyas frecuencias de modulación se sitúan cerca de 70 Hz. Desde el punto de vista perceptivo, se han señalado dos aspectos no totalmente coincidentes según que la frecuencia de la señal portadora sea baja (500 Hz o inferior), o alta (2 KHz o mayor). Este tipo de señales se describen ventajosamente mediante ecuaciones del tipo que sigue:

$$s(t) = A(1 + m \cos 2\pi f_m t) \cos(2\pi f_c t)$$

donde A es la amplitud, m el factor de modulación, f_m la frecuencia de modulación y f_c la frecuencia de la portadora o tono puro.

El factor de modulación aparece como la magnitud clave ligada a la aspereza y para tonos puros modulados en amplitud se admite que la aspereza sigue una función potencial de m cuyo exponente es cercano a 2 [12], [15].

La extracción del parámetro m aparece como una tarea primordial en el cálculo de la aspereza de estas señales. Un algoritmo directo y un tanto ingenuo de extracción lo constituyen la obtención del máximo y del mínimo de estas señales, cuya relación con m viene dado por las ecuaciones siguientes:

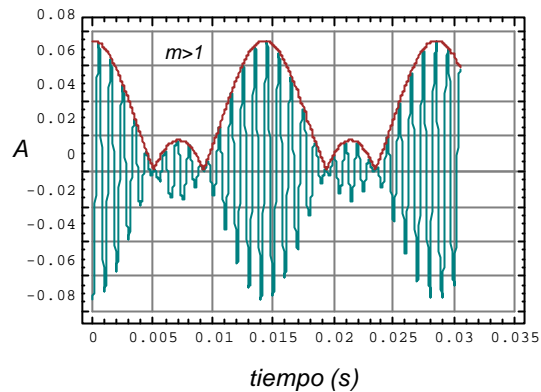
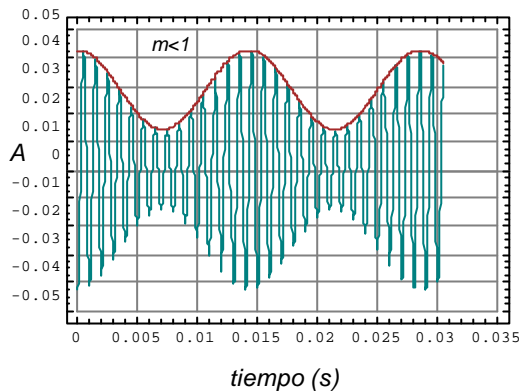
$$\max\{s(t)\} = A(1 + m)$$

$$\min\{s(t)\} = A(1 - m)$$

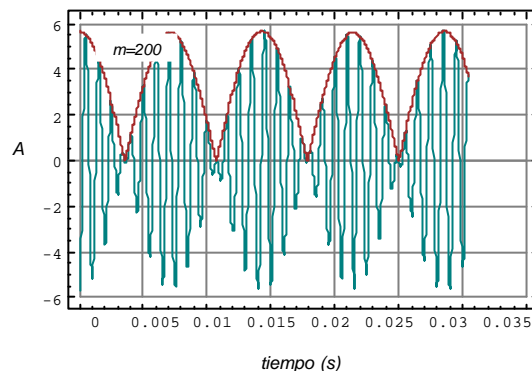
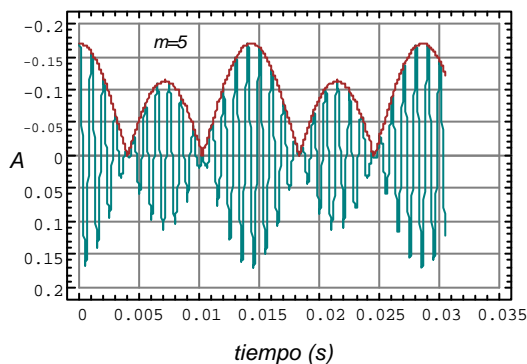
de donde la obtención de m es inmediata. Pero este algoritmo no es válido para $m > 1$, al ser 0 los mínimos. Se resuelve sustituyendo el mínimo por el máximo secundario local que se relaciona con m por la expresión:

$$\max_{\text{sec}}\{s(t)\} = |A(1-m)|$$

Las dos figuras que siguen muestran el efecto de detección de m para las dos situaciones típicas indicadas de $m < 1$ y $m > 1$.



Pero hay otro efecto al crecer m , fácilmente mostrable gráficamente, que se recoge en las figuras siguientes: evolución progresiva al crecer m hacia señales moduladas en amplitud semejantes a simplemente duplicar la frecuencia de modulación.



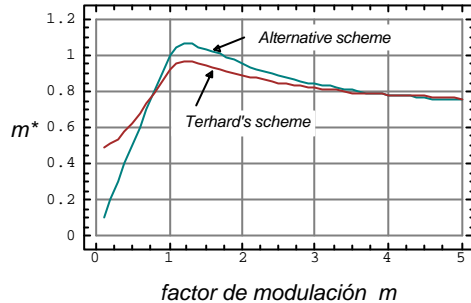
Los resultados subjetivos señalan que la aspereza es una función sensiblemente cuadrática de m para $m < 1$ [15]. Sigue después una función creciente de menor pendiente hasta alcanzar un máximo para $m = 1.2$, disminuyendo después con pendiente decreciente al crecer m [6].

Se observa pues un paralelismo de la modulación efectiva de la amplitud con la intensidad de la percepción subjetiva de la aspereza asociada a estas señales.

La ley cuadrática de dependencia de la aspereza en función de m , es solo válida hasta $m = 1$, decreciendo después en la forma indicada. Esto indica la necesidad de algoritmos que aproximar adecuadamente la ley de variación de la aspereza para cualquier valor m , en las señales moduladas en amplitud y el parámetro correspondiente, por analogía con m , se acostumbra a designar como factor generalizado de modulación y se representa por m^* .

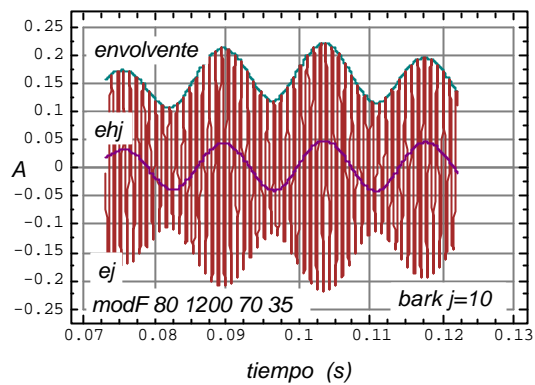
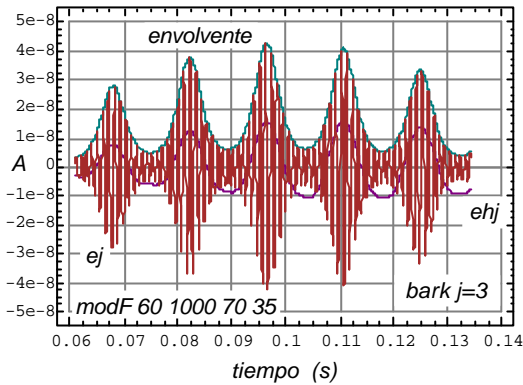
Admitiendo el aserto tan reiterado de que la aspereza está ligada a las fluctuaciones de la envolvente de las señales no es difícil observar que el cociente entre el valor rms de la envolvente de la señal (en sentido matemático estricto) y su valor medio sigue rigurosamente la evolución de la aspereza con m indicada. El esquema operativo de Aures [16], opera de manera análoga pero usando el modulo de la señal en vez de la envolvente. La figura que sigue compara los resultados de ambos procedimientos para un tono puro de 1000 Hz y nivel de 60 dB, modulado por una señal de 70 Hz, cuando m varía desde 0.1 hasta 5. Siguiendo evoluciones de una cierta semejanza se aprecia una más neta

coincidencia de nuestra propuesta con los resultados subjetivos, hasta $m < 3$ tendiendo asintóticamente hacia los mismos valores al crecer m .



Las excitaciones específicas en cada banda de bark, se configuran por superposición de las componentes propias de la banda con las migraciones de cada componente espectral de la señal, migraciones cuya cuantía se evalúa mediante filtros, propuestos por Terhardt [17], con efectos alineales dependientes del nivel y de la frecuencia en la rama descendente hacia altas frecuencias. También Terhardt postula la necesidad de otros filtros suplementarios, que denomina filtros H, para regular dentro de cada banda bark, la contribución de las distintas frecuencias presentes. Son filtros de bajas frecuencias, con una muy lenta evolución al recorrer las 24 bandas de frecuencia subjetiva. La dificultad de determinación subjetiva de estos filtros hace que se hayan obtenido tres claramente distintivos obteniendo los restantes por interpolación de ellos. Esta ponderación de los filtros H, que convierte las excitaciones específicas e_j en las excitaciones filtradas eh_j (que ya son las fluctuaciones de las envolventes) introduce un elemento de distorsión en los esquemas anteriores de obtención de los factores generalizados de modulación ya que operando en e_j no se introduce el necesario efecto de H y operando en eh_j el valor medio temporal es cero. Terhardt propone como operación de cálculo un sistema híbrido: el cociente entre el valor rms de eh_j y el valor medio del módulo de e_j . Nuestro esquema para poder mantener el efecto de los filtros H, también se ha convertido en un sistema híbrido.

Cuando H elimina componentes significativas de e_j las funciones eh_j difieren de las fluctuaciones de la envolvente de e_j . De otro modo son indistinguibles. En la figura que sigue se reproduce un ejemplo de discrepancia que corresponde a la banda $j=3$ barks, de una señal de 1000 Hz, y nivel de 60 dB, modulada en frecuencia con una señal de 70 Hz con una excursión de frecuencia de 35 Hz. El ejemplo de coincidencia corresponde a la banda $j=10$ barks de una portadora de 1200 Hz y nivel de 80 dB, modulada como en el caso anterior por una señal de 70 Hz con una excursión en frecuencia de 35 Hz. (Nótese que la aspereza intrínseca del primer caso es mayor que la del segundo aun cuando la amplitud es del orden de 10^{-7} veces menor, es decir insignificante. También es perceptible en estas figuras, sobre todo en la primera la influencia de la ventana temporal de Blackmann usada, en asociación a la FFT, en el cálculo del espectro).



Los valores de los factores generalizados de modulación calculados con estos dos esquemas difieren en la cuarta cifra significativa, lo que indicaría una posible laxitud en atribuir la aspereza a las fluctuaciones de la envolvente, entendida en sentido matemático estricto.

En señales moduladas en frecuencia encontramos un paralelismo muy grande si se formulan estas señales conforme a la ecuación:

$$s(t) = A \cdot \sin\left(2\pi f_c t - \frac{\Delta f}{f_m} \cos(2\pi f_m t)\right)$$

Los espectros respecto a la señal modulada en amplitud asociada son indistinguibles hasta $m=0.8$. Al aumentar m los espectros de las parejas de señales siguen evoluciones que guardan semejanzas pero se acentúan progresivamente las diferencias, generalmente por incremento diferente de la importancia relativa de los distintos armónicos $f_c \pm n \cdot f_m, n=1,2,..$ dando lugar a señales de excitación que también difieren progresivamente.

Que señales de envolvente plana, como son las señales moduladas en frecuencia, den lugar a percepciones de fluctuación de la envolvente se explica si se atribuye al efecto del filtrado por cada banda crítica, de manera semejante a como se realizan las detecciones de timbre mediante detección de variaciones de amplitud [18]. Pero sucede que dentro de las bandas de bark se observan sistemáticamente auténticas señales moduladas en amplitud como se aprecia en los ejemplos de las dos figuras anteriores que como se ha indicado corresponden a señales moduladas en frecuencia, elegidas simplemente para ilustrar el efecto ya indicado.

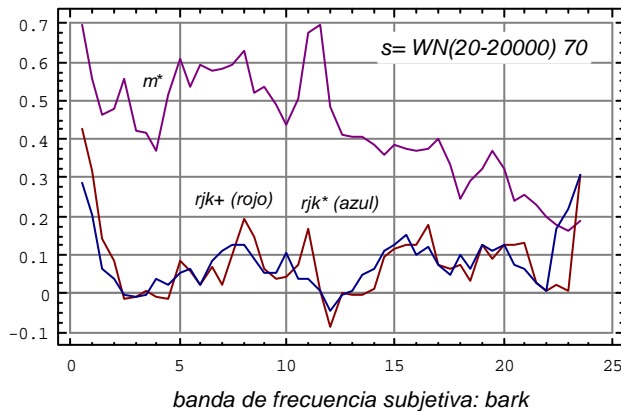
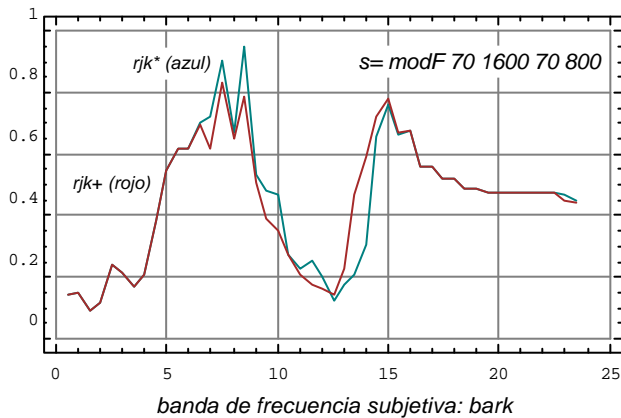
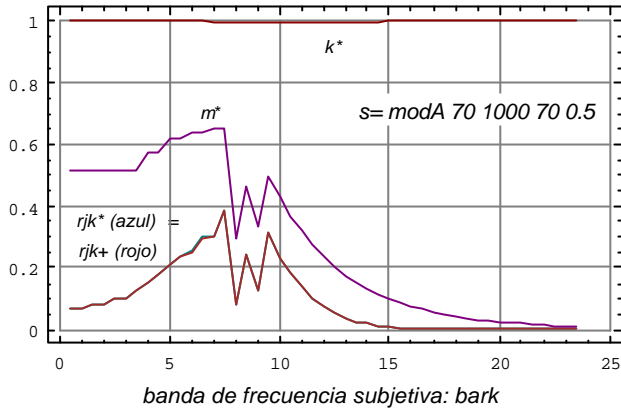
EFFECTOS DE LAS CORRELACIONES ENTRE LAS SEÑALES DE EXCITACIÓN DE LAS BANDAS DE BARK VECINAS EN LA ASPEREZA RESULTANTE

Para reducir la aspereza calculada de ruidos no modulados, que con este modelo resultaba muy por encima de los valores percibidos, Aures propone usar la media de los factores de correlación cruzada de la banda de frecuencia considerada con las vecinas anterior y posterior $(k_{j-1}+k_{j+1})/2$. Daniel y Weber, por su parte usan el cuadrado de su producto $(k_{j-1} \cdot k_{j+1})^2$. Hacemos notar que el efecto de estos factores de correlación se limita prácticamente a las bandas cuyo contenido energético es significativo.

Estos factores de correlación, reducen prácticamente a 0, por doquier, la aspereza específica de ruidos de banda ancha, próximos al ruido blanco. En el caso de señales moduladas en frecuencia los

efectos de estos factores de correlación son patentes solamente en el intervalo $f_c \pm \Delta f$. En señales moduladas en amplitud la aspereza no se ve afectada significativamente por estos términos k . La diferencia de los efectos de tomar la opción de Aures o de D&W no parece en general significativa n el cómputo total de la aspereza.

En las tres figuras que siguen se ilustran estos asertos para los tres tipos de señales considerados.



SIMILITUD ENTRE LAS SEÑALES DE EXCITACIÓN ESPECIFICA ASOCIADAS A SEÑALES MODULADAS EN AMPLITUD: UMBRALES DE ACTUACION EN LA ASPEREZA

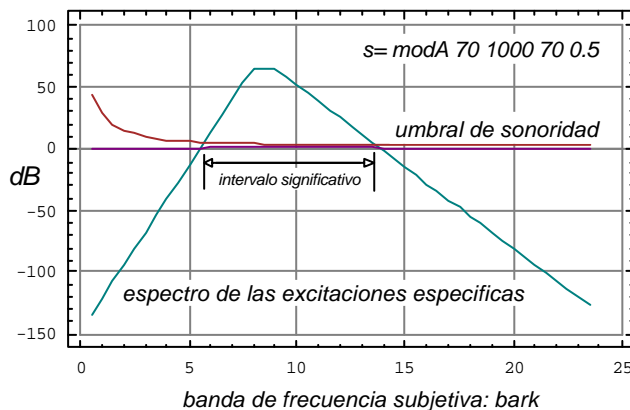
En las bandas de frecuencia críticas alejadas de las bandas de mayor contenido energético, es normal encontrar valores elevados de los factores generalizados de modulación m^* . No obstante su nivel sonoro es muy bajo. Nos encontramos pues en una situación en que todos los valores de aspereza específica r_j , en las distintas bandas, son altos lo que da como resultado valores también elevados de la aspereza total resultante. Urge por tanto encontrar fundamentos para desestimar estos valores, de manera algo semejante a como hemos hecho mediante los factores de correlación k para los ruidos de banda ancha. En un trabajo anterior [2] realizábamos esta operación mediante un operador que excluía las bandas críticas con niveles sonoros 30 dB por debajo de la de mayor contenido energético y obteníamos resultados totales de aspereza casi coincidentes a los dados por D&W para la generalidad de las señales.

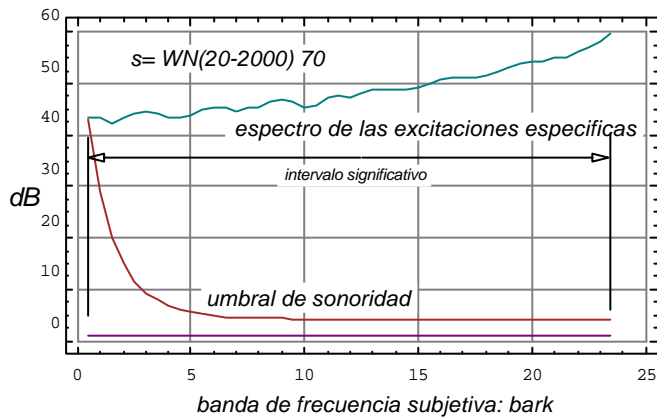
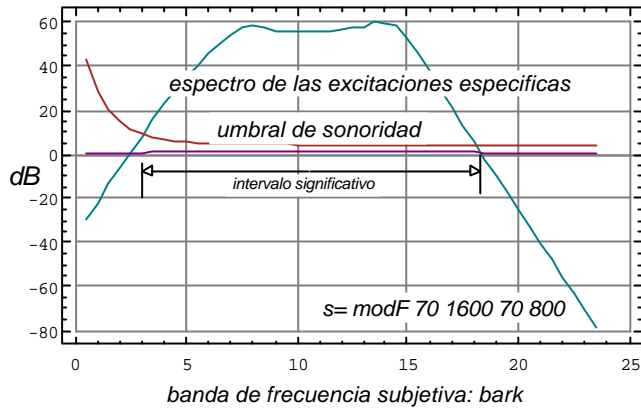
Pero en la literatura de la Psicoacústica se pueden encontrar dos conceptos apropiados a la solución de este problema: el umbral de detección de sonoridad y el umbral diferencial de sonoridad[18]. Siendo la aspereza debida a detecciones de fluctuaciones de la envolvente de la señal la detección de estas variaciones es crucial y es clara que los conceptos enunciados juegan entonces un papel destacada.

Las bandas alejadas de las de mayor contenido energético, que en los tonos puros modulados en amplitud corresponden a las primeras y a las ultimas bandas criticas (bandas con un número de barks inferior a cinco y superiores 20) tienen un nivel muy por debajo del umbral de detección de sonoridad (en ocasiones denominado ruido interno del oído) [19], y por tanto no son percibidas por el oído. Solamente las bandas con niveles de señal de excitación por encima de este umbral deben contribuir a la aspereza.

Semejantemente el umbral diferencial de sonoridad [18], constituye otro motivo de rechazo de las componentes de aspereza en bandas con fluctuaciones de amplitud por debajo de este umbral. Sin embargo hemos comprobado que es siempre mucho menos restrictivo que el anterior y no hace falta tenerse en cuenta.

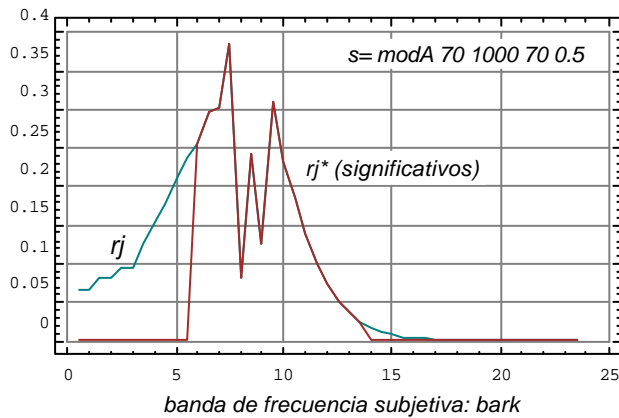
En las tres figuras que siguen se dan ejemplos de las envolventes de las figuras de migración de todas componentes espectrales a las excitaciones específicas, para un tono puro modulado en amplitud, un tono puro modulado en frecuencia y un ruido blanco de banda ancha sintético.

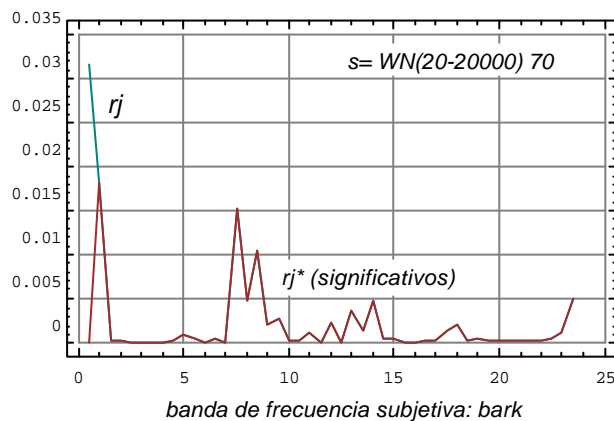
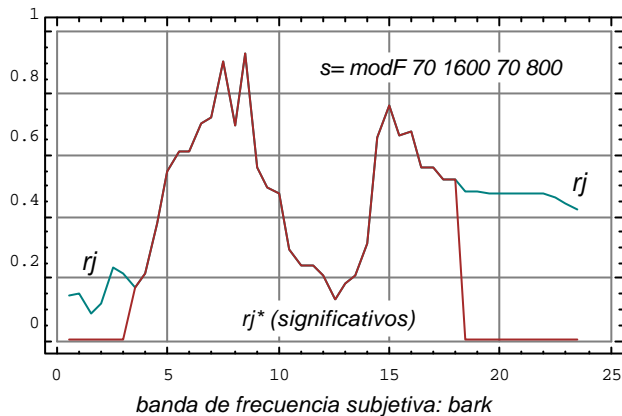




En las mismas gráficas se destacan los intervalos significativos para el cálculo final de la aspereza total resultante.

En las tres gráficas que siguen se dan para las mismas tres señales anteriores las asperezas específicas r_j (espectro de aspereza) significativas y no significativas en la aspereza total resultante.





El tercero de los casos correspondiente a ruido de banda ancha, aproximación del ruido blanco por síntesis directa sencilla, destaca en esta comparación porque todas las bandas críticas son significativas, excepción hecha de la primera, que se debe claramente a deficiencias del procedimiento usado para la síntesis de la señal.

INTERROGANTES Y CONCLUSIONES EN LA PERCEPCION DE LA ASPEREZA DE SEÑALES DEL MEDIO URBANO

La comparación de los resultados de nuestro programa y los valores subjetivos de 32 señales diversas del universo de los ruidos medidos resulta muy divergente [1]. Solamente para los sonidos sintéticos, de entre los anteriores, la correlación es buena. Para casi todos los sonidos, cuyas asperezas calculadas se sitúan entre 0,02 y 4 aproximadamente, las estimaciones quedan en una zona reducida de 1 a 3. No puede afirmarse por tanto que el modelo de cálculo objetivo permita predecir la aspereza de los sonidos del entorno acústico.

Por lo esmerado de las pruebas subjetivas, en si mismas, cabe pensar que la sensación de aspereza quede enmascarada por otros atributos más importantes. Pero más bien creemos en la existencia de otro atributo estrechamente relacionado con la aspereza entendida en el sentido desarrollado en los apartados anteriores que podríamos llamar "académico", y que por tanto no reside en los mismos factores físicos de las señales: otra clase de aspereza con concomitancias y diferencias con el concepto clásico o académico de aspereza. Dejamos la interrogante para motivar a algún estudioso interesado en el tema.

BIBLIOGRAFIA

- [1] G. Canévet et al., Nuevos estudios de validacion subjetiva de los indices de calidad sonora, XXX Jornadas Nacionales de Acúsica, Tecniacústica 99, Avila, 1999
- [2] A. Moreno, F. Simón, C. de la Colina, A. Marchioni, Factores determinantes en el cálculo de la aspereza de señales acústicas por el metodo de Aures, XXX Jornadas Nacionales de Acúsica, Tecniacústica 99, Avila, 1999
- [3] W. Aures, Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, Acustica 59 (1985), 286
- [4] E. Zwicker y E. Terhardt, Facts and models in hearing, Springer Verlag. New York 19174
- [5] EAA Tutorium, Forum Acusticum, Antwerpen 1996
- [6] P. Daniel, R. Weber, Psychoacoustical Roughness: Implementation of an Optimized Model, Acustica 81 (1995)
- [7] H. v. Helmholtz, Die Lehre von den Tonempfindungen als physiologische Grundage für die Theorie der Musik, F. Wieweg & Sohnn Braunsweig, 1863
- [8] G. v. Békésy, Über akustische Rauigkeit, Z. tech. Phy. 16(1935) 276
- [9] E. Terhardt, Pitch, consonance and harmony, J. Acoust. Soc. Amer. 55 (1974), 1061
- [10] W. Aures, Der sensorische Wohlklang als Funktion psycoaakustischer Empfindungsgrossen, Acustica 59 (1985) 282
- [11] S. Kemp, Roughness of frequency modulated tones, Acustica 50 (1982), 126
- [12] A. Vogel, Ein gemeinsames Funktionsschema zur Beschreibung der Lautheit und der Rauigkeit, Tesis, TU München 1974.
- [13] R. Plomp, H. J. M. Steeneken, Interference between two simple tones. J. Acoust. Soc. Amer. 43 (1967) 883
- [14] N. Isshiki et al. Differential diagnosis of hoarseness, Folia Phoniatica 21 (1969) 9-19
- [15] E. Terhardt, Über akustische Rouigkeit und Schwankungsstärke, Acoustica 20,1968), 201
- [16] W. Aures, Ein Berechnungsverfahren der Rauigkeit, Acustica 58 (1985), 268
- [17] E. Terhardt, Calculating virtual pitch, Hearing Research 1 (1979)487-498
- [18] M. C. Botte, G. Canevet, L. Demany, C. Sorin, Psychoacoustique et perception auditive, INSERM/CFA/CNET, Paris 1989
- [19] M. Florentine, S. Buus, An excitation pattenr model for intensity discrimination, J. Acoust. Soc. Am. 70, 1646-1654, 1981

AGRADECIMIENTOS

Parte de este estudio ha sido financiado por el proyecto del Plan Nacional de Investigación Científica y Técnica (I+D), AMB98-1029-C04-01. Otra significativa ayuda se ha tenido del programa general Picasso de cooperación España- Francia, 1998